PTO 06-0062

CY=JA DATE=19981113 KIND=A
PN=10-301954

IMAGE RETRIEVAL METHOD AND IMAGE RETRIEVAL SYSTEM
[Gazou kensaku houhou oyobi gazou kensaku shisutemu]

Teruo Hamano, et al.

UNITED STATES PATENT AND TRADEMARK OFFICE
Washington, D.C. October 2005

Translated by: FLS, Inc.

PUBLICATION COUNTRY	(19):	JP
DOCUMENT NUMBER	(11):	100301954
DOCUMENT KIND	(12):	A [PUBLISHED UNEXAMINED APPLICATION]
PUBLICATION DATE	(43):	19981113
APPLICATION NUMBER	(21):	090112161
APPLICATION DATE	(22):	19970430
INTERNATIONAL CLASSIFICATION	(51):	G 06 F 17/30, 12/00; H 04 N 1/00, 1/32
INVENTORS	(72):	HAMANO, TERUO; KEN, KAJII; NOBORU, SONEHARA
APPLICANT	(71):	NIPPON TELEGRAPH AND TELEPHONE CORPORATION
TITLE	(54):	IMAGE RETRIVAL METHOD AND IMAGE RETRIEVAL SYSTEM
FOREIGN TITLE	(54A):	GAZOU KENSAKU HOUHOU OYOBI GAZOU KENSAKU SHISUTEMU

[Claims] /2*

[Claim 1] With respect to an image retrieval method in which an HTML file retrieval request is made to a retrieval device from a facsimile terminal via a public communication network, in which the retrieval device retrieves an HTML file on the Internet in accordance with the URL information and HTTP, and in which this retrieved file is converted into image data that can be output to a facsimile and then becomes transmitted to the facsimile terminal side,

an image retrieval method characterized by storing URL information and number information, etc. that corresponds to said URL information, by storing data that were utilized for file retrievals and data that were transmitted after being converted to images as cache data, by comparing data in order to determine whether the previously retrieved homepage has been changed or not in response to a file retrieval request, by transmitting stored image data if no change has been made, and by generating and transmitting new image data if a change has been made.

[Claim 2] An image retrieval method of Claim 1 characterized by the above-mentioned data comparison being carried out by: related image files being retrieved, stored, and compared in addition to the HTML files being compared; allowing stored cache data to be the image data that will be transmitted if the above-mentioned stored cache data and the currently retrieved data are the same; and executing image conversion if they are different.

---

* Number in the margin indicates pagination in the foreign text.

[Claim 3] With respect to an image retrieval system in which an HTML file retrieval request is made to a retrieval device from a facsimile terminal via a public communication network, in which the retrieval device retrieves an HTML file on the Internet in accordance with the URL information and HTTP, and in which this retrieved file is converted into image data that can be output to a facsimile and then becomes transmitted to the facsimile terminal side,

an image retrieval system characterized by the above-mentioned retrieval device being equipped with: a line response control part that has a means for controlling line connection responses for the facsimile terminal, analyzing the PB tones input from the facsimile terminal, and outputting image data to a facsimile terminal; a WWW search part that retrieves HTML files on the WWW in accordance with HTTP; a URL address/number converting part that converts numbers into URL addresses based on number information corresponding to accumulated URL information; an image converting part that analyzes HTML files obtained by the above-mentioned WWW search part and that converts them into image data which will be transmitted to the facsimile terminal; a data storing part that stores URL information and number information, etc. corresponding to said URL information; a data cache storing part that stores data retrieved by the above-mentioned WWW search part and the transmitted image data that were converted by the above-mentioned image converting part; and a data comparing part that, regarding the HTML file of a URL address specified by the above-mentioned URL address/number converting part, compares all of the character strings of the previously retrieved HTML

and of the currently retrieved HTML file prior to the execution of image conversion, that utilizes the transmitted image data which is stored in the above-mentioned data cache storing part and which corresponds to said HTML file as an image conversion result if all of the character strings match, and that executes image conversion if the HTML files are different from each other.

[Claim 4] An image retrieval system of Claim 3 characterized by the above-mentioned data comparing part being equipped with a means for retrieving, storing, and comparing related image files in addition to comparing HTML files, for using stored cache data as the image data to be sent if the above-mentioned stored cache data and the currently retrieved data are the same, and for executing image conversion if they are different.

[Detailed Explanation of the Invention]

[0001] [Technical Field of the Invention]

The present invention pertains to image retrieval methods and image retrieval systems for accessing the World Wide Web of the Internet from a facsimile terminal through a network configured by electronic communication equipment and for extracting a retrieved HTML (Hyper Text Markup Language) file after converting it into an image.

[0002] [Related Art]

An example of a system capable of extracting information by accessing WWW of the Internet from a facsimile terminal is a Web Fax (Literature: The Intranet Journal, February 1997, SOFTBANK, pp.180 ~ 181).

[0003] One example of a conventional system is illustrated in Fig. 3. According to this system, a retrieval request is made to a retrieval device [3] from a facsimile terminal [1] via a public communication network [2] and the retrieval device [3] retrieves an HTML file on the Internet, converts this retrieved file into image data that can be output to a facsimile, and transmits it to the facsimile side [1].

[0004] The retrieval device [3] is equipped with: a line response control part that is equipped with a means for controlling line connection responses, analyzing the PB tones input by the user, and outputting the data of a transmitted image to a facsimile terminal; a WWW search part that retrieves data on the WWW in accordance with HTTP (Hyper Text Transformer Protocol); an URL address/number converting part that converts numbers into URL (Uniform Resource Locator) addresses based on accumulated URL information and corresponding number information; and an image converting part that analyzes HTML files obtained by the WWW search part and that converts them into image data to be transmitted to the facsimile terminal.

[0005] Such a system has a problem in that it takes time for data to be transmitted since an HTML file is retrieved and converted into an image every time a URL (Uniform Resource Locator) address is specified from a facsimile terminal by means of push buttons.

[0006] For this reason, there is a cache-type file retrieval system in which retrieval times are shortened by caching HTML files by means of the cache function of the WWW server and by utilizing the cached /3 HTML files when retrieving previously visited homepages.

[0007] [Problem that the Invention is to Solve]

A conventional cache-type file retrieval system has a problem in that, although it can shorten file retrieval times, it cannot shorten image conversion times since image conversion is executed for each file that is retrieved.

[0008] The purpose of the invention is to supply an image retrieval method and an image retrieval system capable of shortening the image conversion times when searching the WWW on the Internet from a facsimile terminal via a network.

[0009] [Means for Solving the Problem]

With respect to an image retrieval method in which an HTML file retrieval request is made to a retrieval device from a facsimile terminal via a public communication network, in which the retrieval device retrieves an HTML file on the Internet in accordance with the URL information and HTTP, and in which this retrieved file is converted into image data that can be output to a facsimile and then becomes transmitted to the facsimile terminal side, the present invention is characterized by storing URL information and number information, etc. that corresponds to said URL information, by storing data that were utilized for file retrievals and data that were transmitted after being converted to images as cache data, by comparing data in order to determine whether the previously retrieved homepage has been changed or not in response to a file retrieval request, by transmitting stored image data if no change has been made, and by generating and transmitting new image data if a change has been made.

[0010] It is also characterized by, in the above-mentioned data comparison, related image files being retrieved, stored, and compared in addition to the HTML files being compared, by allowing stored cache data to be the image data that will be transmitted if the above-mentioned stored cache data and the currently retrieved data are the same, and by executing image conversion if they are different.

[0011] Moreover, with respect to an image retrieval system in which an HTML file retrieval request is made to a retrieval device from a facsimile terminal via a public communication network, in which the retrieval device retrieves an HTML file on the Internet in accordance with the URL information and HTTP, and in which this retrieved file is converted into image data that can be output to a facsimile and then becomes transmitted to the facsimile terminal side, the invention is characterized by the above-mentioned retrieval device being equipped with: a line response control part that has a means for controlling line connection responses for the facsimile terminal, analyzing the PB tones input from the facsimile terminal, and outputting image data to a facsimile terminal; a WWW search part that retrieves HTML files on the WWW in accordance with HTTP; an URL address/number converting part that converts numbers into URL addresses based on number information corresponding to accumulated URL information; an image converting part that analyzes HTML files obtained by the above-mentioned WWW search part and that converts them into image data which will be transmitted to the facsimile terminal; a data storing part that stores URL information and number information, etc., corresponding to said URL information; a data cache storing part that

stores data retrieved by the above-mentioned WWW search part and the transmitted image data that were converted by the above-mentioned image converting part; and a data comparing part that, regarding the HTML file of a URL address specified by the above-mentioned URL address/number converting part, compares all of the character strings of the previously retrieved HTML and of the currently retrieved HTML file prior to the execution of image conversion, that utilizes the transmitted image data which is stored in the above-mentioned data cache storing part and which corresponds to said HTML file as an image conversion result if all of the character strings match, and that executes image conversion if the HTML files are different from each other.

[0012] The process flow of the present invention is indicated in Fig. 1. In the figure, the URL of the homepage to be retrieved is specified by means of PB numbers (Step S1). The URL number is converted into an URL address (Step S2) and a HTML file becomes retrieved (Step S3). The HTML file becomes stored for comparison and caching (Step S4).

[0013] Next, a previously retrieved HTML file that has the same URL address and the currently retrieved HTML file are compared (Step S5). This comparison is carried out based on whether or not all of the character strings within the files are the same.

[0014] If the HTML files are exactly the same, the transmitted image data that has been cached as a result of previous retrieval and image conversion becomes extracted (Step S6) and transmitted to the facsimile terminal (Step S9).

[0015] If the HTML files are different, related image data files are also retrieved and the HTML file becomes converted into an image (Step S7). The generated image data becomes cached for subsequent retrievals (Step S8), and the obtained image data becomes transmitted to the facsimile terminal (Step S9).

[0016] Roughly the same procedure is carried out when related image files are also compared. A difference is that, in addition to the comparison of the HTML file, the related image files are also subjected to binary comparisons at the time of data comparison.

[0017] As for the comparison executed to determine whether the homepage of the specified URL address has been changed or not, the presence or absence of a change cannot be always confirmed accurately by merely comparing the update date and/or size of the specified HTML file. Therefore, (1) an HTML file is retrieved every time, and the previous HTML file and the current HTML file are compared including the character strings within the files. Moreover, (2) related image files as well as the HTML files are subjected to binary data comparisons. /4

[0018] [Embodiment of the Invention]

An embodiment of the present invention is illustrated in Fig. 2. An image retrieval device [10] connected with a facsimile terminal [2] via a public communication network [2] is equipped with retrieval processing components, [11] ~ [16].

[0019] [11] is a line response control part that is equipped with a means for controlling line connection responses, analyzing the PB tones input by the user, and outputting image data to a facsimile terminal.

[12] is an URL address/number converting part that converts numbers into URL (Uniform Resource Locator) addresses based on accumulated URL information and corresponding number information. [13] is a data storing part that stores URL information and number information, etc. corresponding to said URL information. [14] is a WWW search part that searches data on the WWW in accordance with HTTP. [15] is a data cache storing part that stores the data retrieved by the WWW search part [14] and the transmitted image data that was converted by an image converting part [16A].

[0020] [16] is a data comparing part that, regarding the HTML file of a specified URL address, compares all of the character strings of the previously retrieved HTML and of the currently retrieved HTML file prior to image conversion, that determines that the transmitted image data which is stored in the data cache storing part [15] and which corresponds to said HTML file to be the result of image conversion if all of the character strings match, and that executes image conversion if the two HTML files are different from each other. [16A] is the image converting part that analyzes an HTML file obtained by the above-mentioned WWW search part [14] and converts it into image data that will be sent to the facsimile terminal.

[0021] In the above structure, the user operating the facsimile terminal [1] calls the image retrieval device [10] and enters the number corresponding to the URL address of a desired homepage by means of push buttons, etc.

[0022] The input PB tones become converted into a number in the line control part [11] and sent to the URL address/number converting part [12] as a number.

[0023] The URL address/number converting part [12] converts the number into a URL address based on the URL information and the number information corresponding to said URL information stored in the data storing part [13].

[0024] The WWW search part [14] searches the HTML files accumulated in WWW servers, $[20_1]$ ~ $[20_N]$, by using the URL address specified by the URL address/number converting part [12]. A retrieved HTML file becomes stored in the data cache storing part [15].

[0025] After the WWW has been searched, the data are compared in the data comparing part [16]. Concretely speaking, a previously retrieved HTML file corresponding to the currently specified URL is read from the data cache storing part [15] and becomes compared with the currently retrieved HTML file.

[0026] In the comparing method, all of the character strings within the HTML files are compared. At this time, if all of the character strings of the two are the same, image data that is stored in the data cache storing part [15] and that corresponds to said HTML file is sent to the line control part [11] as a result of image conversion instead of actually executing image conversion.

[0027] If the two HTML files are different, the HTML file that was obtained by the WWW search part [14] becomes analyzed in the image converting part [16A], and related files become retrieved.

[0028] Retrieval of the related files may be performed prior to comparison, but the time it takes to retrieve the related files becomes wasted if there is no change in the HTML file.

[0029] If two HTML files are different, image conversion, which is for obtaining image data that will be sent to the facsimile terminal, is carried out, and image data becomes generated and stored in the data cache storing part [15]. Moreover, the stored image data becomes sent to the facsimile device [1] via the line control part [11].

[0030] Moreover, data stored in the data cache storing part [15] can be utilized as cache data in the next search by matching and storing the URL address and the HTML file and/or transmitted image data.

[0031] Moreover, when also comparing related image files, the HTML file and its related files are retrieved prior to the comparison. Then, the data of the related images are compared in addition to the comparison of the character strings of the HTML files in the data comparing part [16]. As for the comparing method, the binary data of the image data are compared. The rest of the process is the same as that mentioned earlier.

[0032] [Effects of the Invention]

As explained earlier, with respect to WWW searches in which a facsimile is utilized, the image retrieval device caches an HTML file and its transmitted image data and the previously cached transmitted image data is sent to a facsimile terminal if there is no change in the homepage of the specified URL address in the present invention. Therefore, if a homepage has been retrieved in the past and if no change has been made to it since then, it is possible to skip the converting process that is

for obtaining image data and to receive image data at high speed.

[0033] In particular, when multiple users are accessing the present device from different facsimile terminals, different users frequently access the same pages, but since image conversions for the facsimiles are not required, it becomes possible to extract data that much faster.

[Brief Explanation of the Drawings]

[Figure 1] A process flow of an image retrieval system of the invention.

[Figure 2] A block diagram showing an embodiment of the invention.

[Figure 3] A WWW search system in which a facsimile terminal is utilized.

[Explanation of the Reference Numerals] /5

[1] = facsimile terminal

[2] = public communication network

[3],[10] = image retrieval device

[11] = line control part

[12] = URL address/number converting part

[13] = data storing part

[14] = WWW search part

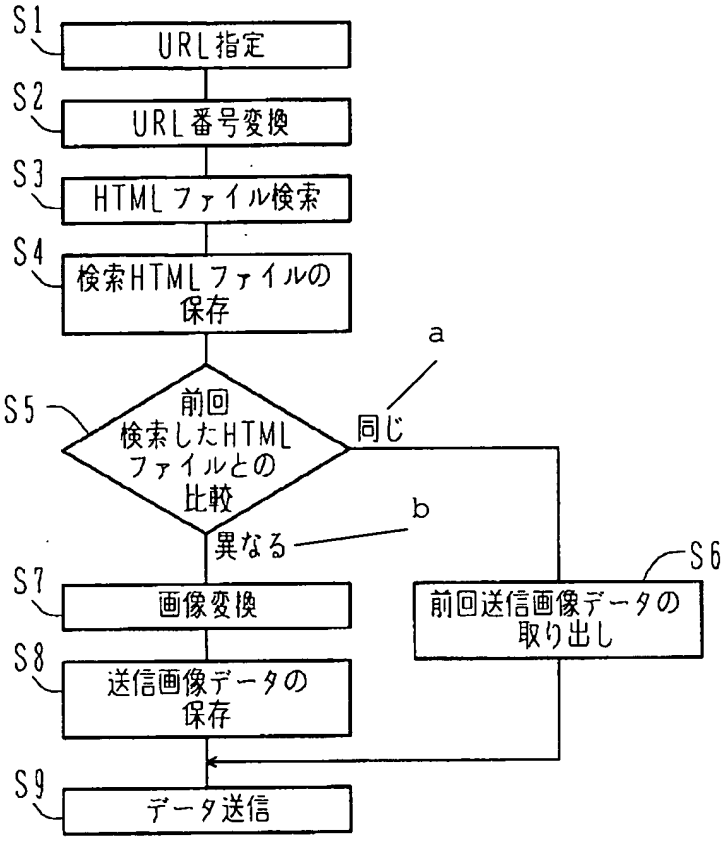
[15] = data cache storing part

[16] = data comparing part

[16A] = image converting part

[Figure 1]

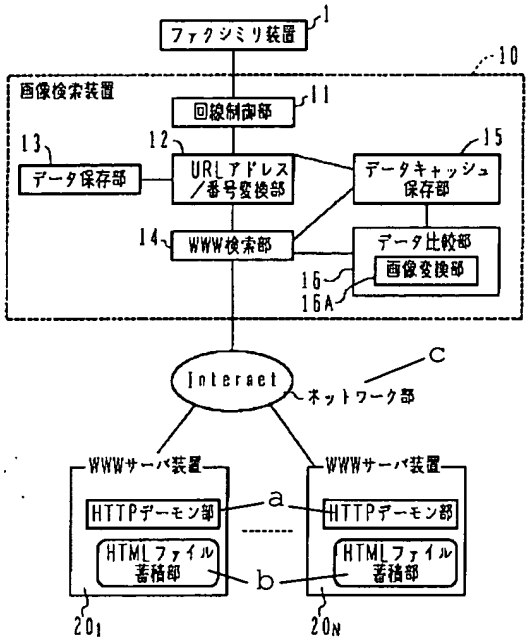
Processing Flow



Key: S1)A URL is specified.; S2)The URL number is converted.; S3)An HTML file is retrieved.; S4)The retrieved HTML file is stored.; S5)A comparison is made with the previously retrieved HTML.; S6)The previously transmitted image data is extracted.; S7)An image is obtained by means of conversion.; S8)The image data is stored.; S9)The data is transmitted.; a)Same; b)Different.

[Figure 2]

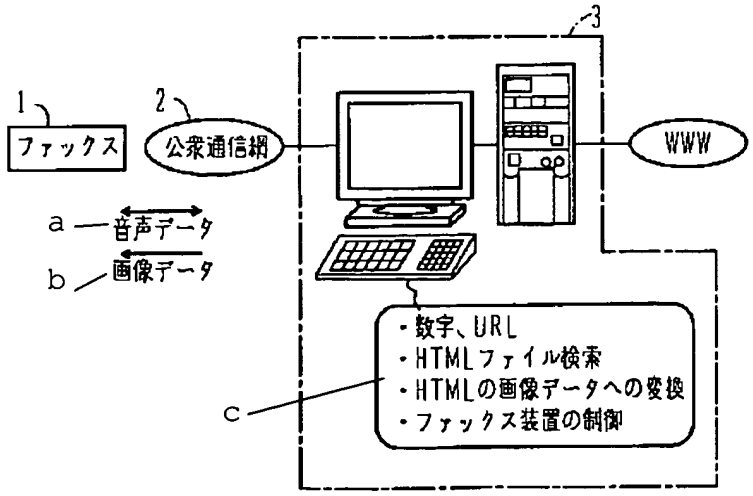
Embodiment



Key: 1) facsimile device; 10) image retrieval device; 11) line control part; 12) URL address/number converting part; 13) data storing part; 14) WWW search part; 15) data cache storing part; 16) data comparing part; 16A) image converting part; 201,20N) WWW server device; a) HTTP daemon part; b) HTML file accumulating part; c) network part.

[Figure 3]

WWW search system with a facsimile terminal



Key: 1) FAX; 2) public communication network; a) voice data; b) image data; c) - Numbers and URL, - HTML file retrieval, - Conversion of HTML into image data, - FAX device control.